



## SEQUENCE LISTING

### (1) GENERAL INFORMATION:

- (i) APPLICANT: E. I. DU PONT DE NEMOURS AND COMPANY
- (ii) TITLE OF INVENTION: CHIMERIC GENES AND METHODS FOR INCREASING THE LYSINE AND THREONINE CONTENT OF THE SEEDS OF PLANTS
- (iii) NUMBER OF SEQUENCES: 108
- (iv) CORRESPONDENCE ADDRESS:
  - (A) ADDRESSEE: E. I. DU PONT DE NEMOURS AND COMPANY
  - (B) STREET: 1007 MARKET STREET
  - (C) CITY: WILMINGTON
  - (D) STATE: DELAWARE
  - (E) COUNTRY: U.S.A.
  - (F) ZIP: 19898
- (v) COMPUTER READABLE FORM:
  - (A) MEDIUM TYPE: FLOPPY DISK
  - (B) COMPUTER: IBM PC COMPATIBLE
  - (C) OPERATING SYSTEM: PC-DOS/MS-DOS
  - (D) SOFTWARE: MICROSOFT WORD VERSION 2.0C
- (vi) CURRENT APPLICATION DATA:
  - (A) APPLICATION NUMBER: US/10/023,066
  - (B) FILING DATE: 17-Dec-2001
  - (C) CLASSIFICATION:
- (viii) ATTORNEY/AGENT INFORMATION:
  - (A) NAME: BARBARA C. SIEGELL
  - (B) REGISTRATION NUMBER: 30,684
  - (C) REFERENCE/DOCKET NUMBER: BB-1037-C
- (ix) TELECOMMUNICATION INFORMATION:
  - (A) TELEPHONE: 302-992-4931
  - (B) TELEFAX: 302-773-0164
  - (C) TELEX: 835420

### (2) INFORMATION FOR SEQ ID NO:1:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 1350 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- (ix) FEATURE:
  - (A) NAME/KEY: CDS
  - (B) LOCATION: 1..1350
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

ATG	GCT	GAA	ATT	GTT	GTC	TCC	AAA	TTT	GGC	GGT	ACC	AGC	GTA	GCT	GAT	48
Met	Ala	Glu	Ile	Val	Val	Ser	Lys	Phe	Gly	Gly	Thr	Ser	Val	Ala	Asp	
1				5					10					15		
TTT	GAC	GCC	ATG	AAC	CGC	AGC	GCT	GAT	ATT	GTG	CTT	TCT	GAT	GCC	AAC	96
Phe	Asp	Ala	Met	Asn	Arg	Ser	Ala	Asp	Ile	Val	Leu	Ser	Asp	Ala	Asn	
			20					25					30			
GTG	CGT	TTA	GTT	GTC	CTC	TCG	GCT	TCT	GCT	GGT	ATC	ACT	AAT	CTG	CTG	144
Val	Arg	Leu	Val	Val	Leu	Ser	Ala	Ser	Ala	Gly	Ile	Thr	Asn	Leu	Leu	
		35					40					45				
GTC	GCT	TTA	GCT	GAA	GGA	CTG	GAA	CCT	GGC	GAG	CGA	TTC	GAA	AAA	CTC	192
Val	Ala	Leu	Ala	Glu	Gly	Leu	Glu	Pro	Gly	Glu	Arg	Phe	Glu	Lys	Leu	
	50					55					60					
GAC	GCT	ATC	CGC	AAC	ATC	CAG	TTT	GCC	ATT	CTG	GAA	CGT	CTG	CGT	TAC	240
Asp	Ala	Ile	Arg	Asn	Ile	Gln	Phe	Ala	Ile	Leu	Glu	Arg	Leu	Arg	Tyr	
65					70					75					80	
CCG	AAC	GTT	ATC	CGT	GAA	GAG	ATT	GAA	CGT	CTG	CTG	GAG	AAC	ATT	ACT	288
Pro	Asn	Val	Ile	Arg	Glu	Glu	Ile	Glu	Arg	Leu	Leu	Glu	Asn	Ile	Thr	
				85					90					95		
GTT	CTG	GCA	GAA	GCG	GCG	GCG	CTG	GCA	ACG	TCT	CCG	GCG	CTG	ACA	GAT	336
Val	Leu	Ala	Glu	Ala	Ala	Ala	Leu	Ala	Thr	Ser	Pro	Ala	Leu	Thr	Asp	
			100				105						110			
GAG	CTG	GTC	AGC	CAC	GGC	GAG	CTG	ATG	TCG	ACC	CTG	CTG	TTT	GTT	GAG	384
Glu	Leu	Val	Ser	His	Gly	Glu	Leu	Met	Ser	Thr	Leu	Leu	Phe	Val	Glu	
		115					120					125				
ATC	CTG	CGC	GAA	CGC	GAT	GTT	CAG	GCA	CAG	TGG	TTT	GAT	GTA	CGT	AAA	432
Ile	Leu	Arg	Glu	Arg	Asp	Val	Gln	Ala	Gln	Trp	Phe	Asp	Val	Arg	Lys	
	130					135					140					
GTG	ATG	CGT	ACC	AAC	GAC	CGA	TTT	GGT	CGT	GCA	GAG	CCA	GAT	ATA	GCC	480
Val	Met	Arg	Thr	Asn	Asp	Arg	Phe	Gly	Arg	Ala	Glu	Pro	Asp	Ile	Ala	
145					150					155					160	
GCG	CTG	GCG	GAA	CTG	GCC	GCG	CTG	CAG	CTG	CTC	CCA	CGT	CTC	AAT	GAA	528
Ala	Leu	Ala	Glu	Leu	Ala	Ala	Leu	Gln	Leu	Leu	Pro	Arg	Leu	Asn	Glu	
			165				170							175		
GGC	TTA	GTG	ATC	ACC	CAG	GGA	TTT	ATC	GGT	AGC	GAA	AAT	AAA	GGT	CGT	576
Gly	Leu	Val	Ile	Thr	Gln	Gly	Phe	Ile	Gly	Ser	Glu	Asn	Lys	Gly	Arg	
		180					185						190			
ACA	ACG	ACG	CTT	GGC	CGT	GGA	GGC	AGC	GAT	TAT	ACG	GCA	GCC	TTG	CTG	624
Thr	Thr	Thr	Leu	Gly	Arg	Gly	Gly	Ser	Asp	Tyr	Thr	Ala	Ala	Leu	Leu	
		195				200						205				
GCG	GAG	GCT	TTA	CAC	GCA	TCT	CGT	GTT	GAT	ATC	TGG	ACC	GAC	GTC	CCG	672
Ala	Glu	Ala	Leu	His	Ala	Ser	Arg	Val	Asp	Ile	Trp	Thr	Asp	Val	Pro	
	210					215					220					
GGC	ATC	TAC	ACC	ACC	GAT	CCA	CGC	GTA	GTT	TCC	GCA	GCA	AAA	CGC	ATT	720
Gly	Ile	Tyr	Thr	Thr	Asp	Pro	Arg	Val	Val	Ser	Ala	Ala	Lys	Arg	Ile	
225					230					235					240	

GAT	GAA	ATC	GCG	TTT	GCC	GAA	GCG	GCA	GAG	ATG	GCA	ACT	TTT	GGT	GCA	768
Asp	Glu	Ile	Ala	Phe	Ala	Glu	Ala	Ala	Glu	Met	Ala	Thr	Phe	Gly	Ala	
				245					250					255		
AAA	GTA	CTG	CAT	CCG	GCA	ACG	TTG	CTA	CCC	GCA	GTA	CGC	AGC	GAT	ATC	816
Lys	Val	Leu	His	Pro	Ala	Thr	Leu	Leu	Pro	Ala	Val	Arg	Ser	Asp	Ile	
			260				265					270				
CCG	GTC	TTT	GTC	GGC	TCC	AGC	AAA	GAC	CCA	CGC	GCA	GGT	GGT	ACG	CTG	864
Pro	Val	Phe	Val	Gly	Ser	Ser	Lys	Asp	Pro	Arg	Ala	Gly	Gly	Thr	Leu	
		275					280					285				
GTG	TGC	AAT	AAA	ACT	GAA	AAT	CCG	CCG	CTG	TTC	CGC	GCT	CTG	GCG	CTT	912
Val	Cys	Asn	Lys	Thr	Glu	Asn	Pro	Pro	Leu	Phe	Arg	Ala	Leu	Ala	Leu	
	290					295					300					
CGT	CGC	AAT	CAG	ACT	CTG	CTC	ACT	TTG	CAC	AGC	CTG	AAT	ATG	CTG	CAT	960
Arg	Arg	Asn	Gln	Thr	Leu	Leu	Thr	Leu	His	Ser	Leu	Asn	Met	Leu	His	
305					310					315				320		
TCT	CGC	GGT	TTC	CTC	GCG	GAA	GTT	TTC	GGC	ATC	CTC	GCG	CGG	CAT	AAT	1008
Ser	Arg	Gly	Phe	Leu	Ala	Glu	Val	Phe	Gly	Ile	Leu	Ala	Arg	His	Asn	
			325					330						335		
ATT	TCG	GTA	GAC	TTA	ATC	ACC	ACG	TCA	GAA	GTG	AGC	GTG	GCA	TTA	ACC	1056
Ile	Ser	Val	Asp	Leu	Ile	Thr	Thr	Ser	Glu	Val	Ser	Val	Ala	Leu	Thr	
			340					345					350			
CTT	GAT	ACC	ACC	GGT	TCA	ACC	TCC	ACT	GGC	GAT	ACG	TTG	CTG	ACG	CAA	1104
Leu	Asp	Thr	Thr	Gly	Ser	Thr	Ser	Thr	Gly	Asp	Thr	Leu	Leu	Thr	Gln	
		355					360					365				
TCT	CTG	CTG	ATG	GAG	CTT	TCC	GCA	CTG	TGT	CGG	GTG	GAG	GTG	GAA	GAA	1152
Ser	Leu	Leu	Met	Glu	Leu	Ser	Ala	Leu	Cys	Arg	Val	Glu	Val	Glu	Glu	
	370					375					380					
GGT	CTG	GCG	CTG	GTC	GCG	TTG	ATT	GGC	AAT	GAC	CTG	TCA	AAA	GCC	TGC	1200
Gly	Leu	Ala	Leu	Val	Ala	Leu	Ile	Gly	Asn	Asp	Leu	Ser	Lys	Ala	Cys	
385				390					395					400		
GCC	GTT	GGC	AAA	GAG	GTA	TTC	GGC	GTA	CTG	GAA	CCG	TTC	AAC	ATT	CGC	1248
Ala	Val	Gly	Lys	Glu	Val	Phe	Gly	Val	Leu	Glu	Pro	Phe	Asn	Ile	Arg	
			405				410						415			
ATG	ATT	TGT	TAT	GGC	GCA	TCC	AGC	CAT	AAC	CTG	TGC	TTC	CTG	GTG	CCC	1296
Met	Ile	Cys	Tyr	Gly	Ala	Ser	Ser	His	Asn	Leu	Cys	Phe	Leu	Val	Pro	
		420					425					430				
GGC	GAA	GAT	GCC	GAG	CAG	GTG	GTG	CAA	AAA	CTG	CAT	AGT	AAT	TTG	TTT	1344
Gly	Glu	Asp	Ala	Glu	Gln	Val	Val	Gln	Lys	Leu	His	Ser	Asn	Leu	Phe	
		435				440					445					
GAG	TAA															1350
Glu	*															
	450															

(2) INFORMATION FOR SEQ ID NO:2:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 36 base pairs

- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

GATCCATGGC TGAAATTGTT GTCTCCAAAT TTGGCG 36

(2) INFORMATION FOR SEQ ID NO:3:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 36 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

GTACCGCCAA ATTTGGAGAC AACAATTTCA GCCATG 36

(2) INFORMATION FOR SEQ ID NO:4:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 48 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

CCCGGGCCAT GGCTACAGGT TTAACAGCTA AGACCGGAGT AGAGCACT 48

(2) INFORMATION FOR SEQ ID NO:5:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 37 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

GATATCGAAT TCTCATTATA GAACTCCAGC TTTTTC 37

(2) INFORMATION FOR SEQ ID NO:6:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 917 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

(A) NAME/KEY: CDS

(B) LOCATION: 3..911

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

CC ATG GCT ACA GGT TTA ACA GCT AAG ACC GGA GTA GAG CAC TTC GGC	47
Met Ala Thr Gly Leu Thr Ala Lys Thr Gly Val Glu His Phe Gly	
1 5 10 15	
ACC GTT GGA GTA GCA ATG GTT ACT CCA TTC ACG GAA TCC GGA GAC ATC	95
Thr Val Gly Val Ala Met Val Thr Pro Phe Thr Glu Ser Gly Asp Ile	
20 25 30	
GAT ATC GCT GCT GGC CGC GAA GTC GCG GCT TAT TTG GTT GAT AAG GGC	143
Asp Ile Ala Ala Gly Arg Glu Val Ala Ala Tyr Leu Val Asp Lys Gly	
35 40 45	
TTG GAT TCT TTG GTT CTC GCG GGC ACC ACT GGT GAA TCC CCA ACG ACA	191
Leu Asp Ser Leu Val Leu Ala Gly Thr Thr Gly Glu Ser Pro Thr Thr	
50 55 60	
ACC GCC GCT GAA AAA CTA GAA CTG CTC AAG GCC GTT CGT GAG GAA GTT	239
Thr Ala Ala Glu Lys Leu Glu Leu Leu Lys Ala Val Arg Glu Glu Val	
65 70 75	
GGG GAT CGG GCG AAG CTC ATC GCC GGT GTC GGA ACC AAC AAC ACG CGG	287
Gly Asp Arg Ala Lys Leu Ile Ala Gly Val Gly Thr Asn Asn Thr Arg	
80 85 90 95	
ACA TCT GTG GAA CTT GCG GAA GCT GCT GCT TCT GCT GGC GCA GAC GGC	335
Thr Ser Val Glu Leu Ala Glu Ala Ala Ser Ala Gly Ala Asp Gly	
100 105 110	
CTT TTA GTT GTA ACT CCT TAT TAC TCC AAG CCG AGC CAA GAG GGA TTG	383
Leu Leu Val Val Thr Pro Tyr Tyr Ser Lys Pro Ser Gln Glu Gly Leu	
115 120 125	
CTG GCG CAC TTC GGT GCA ATT GCT GCA GCA ACA GAG GTT CCA ATT TGT	431
Leu Ala His Phe Gly Ala Ile Ala Ala Ala Thr Glu Val Pro Ile Cys	
130 135 140	
CTC TAT GAC ATT CCT GGT CGG TCA GGT ATT CCA ATT GAG TCT GAT ACC	479
Leu Tyr Asp Ile Pro Gly Arg Ser Gly Ile Pro Ile Glu Ser Asp Thr	
145 150 155	
ATG AGA CGC CTG AGT GAA TTA CCT ACG ATT TTG GCG GTC AAG GAC GCC	527
Met Arg Arg Leu Ser Glu Leu Pro Thr Ile Leu Ala Val Lys Asp Ala	
160 165 170 175	
AAG GGT GAC CTC GTT GCA GCC ACG TCA TTG ATC AAA GAA ACG GGA CTT	575
Lys Gly Asp Leu Val Ala Ala Thr Ser Leu Ile Lys Glu Thr Gly Leu	
180 185 190	
GCC TGG TAT TCA GGC GAT GAC CCA CTA AAC CTT GTT TGG CTT GCT TTG	623
Ala Trp Tyr Ser Gly Asp Asp Pro Leu Asn Leu Val Trp Leu Ala Leu	
195 200 205	

GGC GGA TCA GGT TTC ATT TCC GTA ATT GGA CAT GCA GCC CCC ACA GCA 671  
Gly Gly Ser Gly Phe Ile Ser Val Ile Gly His Ala Ala Pro Thr Ala  
210 215 220

TTA CGT GAG TTG TAC ACA AGC TTC GAG GAA GGC GAC CTC GTC CGT GCG 719  
Leu Arg Glu Leu Tyr Thr Ser Phe Glu Glu Gly Asp Leu Val Arg Ala  
225 230 235

CGG GAA ATC AAC GCC AAA CTA TCA CCG CTG GTA GCT GCC CAA GGT CGC 767  
Arg Glu Ile Asn Ala Lys Leu Ser Pro Leu Val Ala Ala Gln Gly Arg  
240 245 250 255

TTG GGT GGA GTC AGC TTG GCA AAA GCT GCT CTG CGT CTG CAG GGC ATC 815  
Leu Gly Gly Val Ser Leu Ala Lys Ala Ala Leu Arg Leu Gln Gly Ile  
260 265 270

AAC GTA GGA GAT CCT CGA CTT CCA ATT ATG GCT CCA AAT GAG CAG GAA 863  
Asn Val Gly Asp Pro Arg Leu Pro Ile Met Ala Pro Asn Glu Gln Glu  
275 280 285

CTT GAG GCT CTC CGA GAA GAC ATG AAA AAA GCT GGA GTT CTA TAA TGAGAATTC 917  
Leu Glu Ala Leu Arg Glu Asp Met Lys Lys Ala Gly Val Leu \*  
290 295 300

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 22 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

CTTCCCGTGA CCATGGGCCA TC

22

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 75 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

CATGGCTGGC TTCCCCACGA GGAAGACCAA CAATGACATT ACCTCCATTG CTAGCAACGG 60

TGGAAGAGTA CAATG

75

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 75 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

CATGCATTGT ACTCTTCCAC CGTTGCTAGC AATGGAGGTA ATGTCATTGT TGGTCTTCCT 60  
CGTGGGGAAG CCAGC 75

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 90 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

CATGGCTTCC TCAATGATCT CCTCCCCAGC TGTTACCACC GTCAACCGTG CCGGTGCCCG 60  
CATGGTTGCT CCATTCACCG GCCTCAAAAG 90

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 90 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

CATGCTTTTG AGGCCGGTGA ATGGAGCAAC CATGCCGGCA CCGGCACGGT TGACGGTGGT 60  
AACAGCTGGG GAGGAGATCA TTGAGGAAGC 90

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 23 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

CCGGTTTGCT GTAATAGGTA CCA 23

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 31 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

AGCTTGGTAC CTATTACAGC AAACCGGCAT G 31

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 27 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

GCTTCCTCAA TGATCTCCTC CCCAGCT 27

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 28 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

CATTGTACTC TTCCACCGTT GCTAGCAA 28

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 20 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:  
 (A) NAME/KEY: misc\_feature  
 (B) LOCATION: 1..20  
 (D) OTHER INFORMATION: /product= "synthetic  
 oligonucleotide"  
 /standard\_name= "SM  
 70"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:



CTGACTCGCT GCGCTCGGTC

20

(2) INFORMATION FOR SEQ ID NO:17:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 24 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- (ix) FEATURE:
  - (A) NAME/KEY: misc\_feature
  - (B) LOCATION: 1..24
  - (D) OTHER INFORMATION: /product= "synthetic oligonucleotide"  
/standard\_name= "SM  
71"
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

TATTTTCTCC TTACGCATCT GTGC

24

(2) INFORMATION FOR SEQ ID NO:18:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 27 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- (ix) FEATURE:
  - (A) NAME/KEY: misc\_feature
  - (B) LOCATION: 1..27
  - (D) OTHER INFORMATION: /product= "synthetic oligonucleotide"  
/standard\_name= "SM  
78"
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

TTCATCGATA GGCGACCACA CCCGTCC

27

(2) INFORMATION FOR SEQ ID NO:19:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 27 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- (ix) FEATURE:
  - (A) NAME/KEY: misc\_feature
  - (B) LOCATION: 1..27
  - (D) OTHER INFORMATION: /product= "synthetic

oligonucleotide"  
/standard\_name= "SM  
79"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

AATATCGATG CCACGATGCG TCCGGCG

27

(2) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 55 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

(A) NAME/KEY: misc\_feature

(B) LOCATION: 1..55

(D) OTHER INFORMATION: /product= "synthetic  
oligonucleotide"  
/standard\_name= "SM  
81"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

CATGGAGGAG AAGATGAAGG CGATGGAAGA GAAGATGAAG GCGTGATAGG TACCG

55

(2) INFORMATION FOR SEQ ID NO:21:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 55 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

(A) NAME/KEY: misc\_feature

(B) LOCATION: 1..55

(D) OTHER INFORMATION: /product= "synthetic  
oligonucleotide"  
/standard\_name= "SM  
80"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

AATTCGGTAC CTATCAGGCC TTCATCTTCT CTTCCATCGC CTTTCATCTTC TCCTC

55

(2) INFORMATION FOR SEQ ID NO:22:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 14 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: unknown

(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(ix) FEATURE:

(A) NAME/KEY: Protein

(B) LOCATION: 1..14

(D) OTHER INFORMATION: /label= name

/note= "base gene

[(SSP5)2]"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

Met Glu Glu Lys Met Lys Ala Met Glu Glu Lys Met Lys Ala  
1 5 10

(2) INFORMATION FOR SEQ ID NO:23:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 21 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

(A) NAME/KEY: misc\_feature

(B) LOCATION: 1..21

(D) OTHER INFORMATION: /product=

"synthetic

oligonucleotide"

/standard\_name= "SM

84"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

GATGGAGGAG AAGATGAAGG C

21

(2) INFORMATION FOR SEQ ID NO:24:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 21 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

(A) NAME/KEY: misc\_feature

(B) LOCATION: 1..21

(D) OTHER INFORMATION: /product= "synthetic

oligonucleotide"

/standard\_name= "SM

85"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

ATCGCCTTCA TCTTCTCCTC C

21

(2) INFORMATION FOR SEQ ID NO:25:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 21 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:  
 (A) NAME/KEY: misc\_feature  
 (B) LOCATION: 1..21  
 (D) OTHER INFORMATION: /product= "synthetic  
 oligonucleotide"  
 /standard\_name= "SM  
 82"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

GATGGAGGAG AAGCTGAAGG C

21

(2) INFORMATION FOR SEQ ID NO:26:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 21 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:  
 (A) NAME/KEY: misc\_feature  
 (B) LOCATION: 1..21  
 (D) OTHER INFORMATION: /product= "synthetic  
 oligonucleotide"  
 /standard\_name= "SM  
 83"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

ATCGCCTTCA GCTTCTCCTC C

21

(2) INFORMATION FOR SEQ ID NO:27:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 7 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: unknown  
 (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

Met Glu Glu Lys Leu Lys Ala  
 1 5

(2) INFORMATION FOR SEQ ID NO:28:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 7 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS: unknown
  - (D) TOPOLOGY: unknown
- (ii) MOLECULE TYPE: protein
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

Met Glu Glu Lys Met Lys Ala  
 1 5

(2) INFORMATION FOR SEQ ID NO:29:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 160 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- (vi) ORIGINAL SOURCE:
  - (B) STRAIN: E. coli
  - (G) CELL TYPE: DH5 alpha
- (vii) IMMEDIATE SOURCE:
  - (B) CLONE: C15
- (ix) FEATURE:
  - (A) NAME/KEY: CDS
  - (B) LOCATION: 2..151
  - (D) OTHER INFORMATION: /function= "synthetic storage protein"  
 /product= "protein"  
 /gene= "ssp"  
 /standard\_name= "5.7.7.7.7.5"
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

```

C ATG GAG GAG AAG ATG AAG GCG ATG GAG GAG AAG CTG AAG GCG ATG   46
  Met Glu Glu Lys Met Lys Ala Met Glu Glu Lys Leu Lys Ala Met
    1           5           10           15

GAG GAG AAG CTG AAG GCG ATG GAG GAG AAG CTG AAG GCG ATG GAG GAG   94
Glu Glu Lys Leu Lys Ala Met Glu Glu Lys Leu Lys Ala Met Glu Glu
          20           25           30

AAG CTG AAG GCG ATG GAG GAG AAG CTG AAG GCG ATG GAA GAG AAG ATG 142
Lys Leu Lys Ala Met Glu Glu Lys Leu Lys Ala Met Glu Glu Lys Met
          35           40           45

AAG GCG TGATAGGTAC CG                                           160
Lys Ala
    50
  
```

(2) INFORMATION FOR SEQ ID NO:30:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 49 amino acids  
(B) TYPE: amino acid  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

Met Glu Glu Lys Met Lys Ala Met Glu Glu Lys Leu Lys Ala Met Glu  
1 5 10 15  
Glu Lys Leu Lys Ala Met Glu Glu Lys Leu Lys Ala Met Glu Glu Lys  
20 25 30  
Leu Lys Ala Met Glu Glu Lys Leu Lys Ala Met Glu Glu Lys Met Lys  
35 40 45  
Ala

(2) INFORMATION FOR SEQ ID NO:31:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 160 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:

- (B) STRAIN: E. coli  
(G) CELL TYPE: DH5 alpha

(vii) IMMEDIATE SOURCE:

- (B) CLONE: C20

(ix) FEATURE:

- (A) NAME/KEY: CDS  
(B) LOCATION: 2..151  
(D) OTHER INFORMATION: /function= "synthetic  
storage protein"  
/product= "protein"  
/gene= "ssp"  
/standard\_name=  
"5.7.7.7.7.5"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

C ATG GAG GAG AAG ATG AAG GCG ATG GAG GAG AAG CTG AAG GCG ATG 46  
Met Glu Glu Lys Met Lys Ala Met Glu Glu Lys Leu Lys Ala Met  
1 5 10 15  
GAG GAG AAG CTG AAG GCG ATG GAG GAG AAG CTG AAG GCG ATG GAG GAG 94  
Glu Glu Lys Leu Lys Ala Met Glu Glu Lys Leu Lys Ala Met Glu Glu  
20 25 30  
AAG CTG AAG GCG ATG GAG GAG AAG CTG AAG GCG ATG GAA GAG AAG ATG 142  
Lys Leu Lys Ala Met Glu Glu Lys Leu Lys Ala Met Glu Glu Lys Met  
35 40 45

AAG GCG TGATAGGTAC CG  
Lys Ala

160

50

(2) INFORMATION FOR SEQ ID NO:32:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 49 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:

Met Glu Glu Lys Met Lys Ala Met Glu Glu Lys Leu Lys Ala Met Glu  
1 5 10 15  
Glu Lys Leu Lys Ala Met Glu Glu Lys Leu Lys Ala Met Glu Glu Lys  
20 25 30  
Leu Lys Ala Met Glu Glu Lys Leu Lys Ala Met Glu Glu Lys Met Lys  
35 40 45

Ala

(2) INFORMATION FOR SEQ ID NO:33:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 139 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:

- (B) STRAIN: E. coli
- (G) CELL TYPE: DH5 alpha

(vii) IMMEDIATE SOURCE:

- (B) CLONE: C30

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 2..130
- (D) OTHER INFORMATION: /function= "synthetic storage protein"  
/product= "protein"  
/gene= "ssp"  
/standard\_name= "5.7.7.7.5"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:

C ATG GAG GAG AAG ATG AAG GCG ATG GAG GAG AAG CTG AAG GCG ATG  
Met Glu Glu Lys Met Lys Ala Met Glu Glu Lys Leu Lys Ala Met  
1 5 10 15

46

GAG GAG AAG CTG AAG GCG ATG GAG GAG AAG CTG AAG GCG ATG GAG GAG 94  
 Glu Glu Lys Leu Lys Ala Met Glu Glu Lys Leu Lys Ala Met Glu Glu  
                   20                  25                  30

AAG CTG AAG GCG ATG GAA GAG AAG ATG AAG GCG TGATAGGTAC CG 139  
 Lys Leu Lys Ala Met Glu Glu Lys Met Lys Ala  
                   35                  40

(2) INFORMATION FOR SEQ ID NO:34:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 42 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:34:

Met Glu Glu Lys Met Lys Ala Met Glu Glu Lys Leu Lys Ala Met Glu  
   1                  5                  10                  15

Glu Lys Leu Lys Ala Met Glu Glu Lys Leu Lys Ala Met Glu Glu Lys  
                   20                  25                  30

Leu Lys Ala Met Glu Glu Lys Met Lys Ala  
                   35                  40

(2) INFORMATION FOR SEQ ID NO:35:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 97 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:

(B) STRAIN: E. coli

(G) CELL TYPE: DH5 alpha

(vii) IMMEDIATE SOURCE:

(B) CLONE: D16

(ix) FEATURE:

(A) NAME/KEY: CDS

(B) LOCATION: 2..88

(D) OTHER INFORMATION: /function= "synthetic  
 storage protein"  
 /product= "protein"  
 /gene= "ssp"  
 /standard\_name=  
 "5.5.5.5"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:35:

C ATG GAG GAG AAG ATG AAG GCG ATG GAG GAG AAG ATG AAG GCG ATG 46  
 Met Glu Glu Lys Met Lys Ala Met Glu Glu Lys Met Lys Ala Met  
   1                  5                  10                  15



GAG GAG AAG ATG AAG GCG ATG GAA GAG AAG ATG AAG GCG TGATAGGTAC 95  
 Glu Glu Lys Met Lys Ala Met Glu Glu Lys Met Lys Ala  
                     20                    25

CG

97

(2) INFORMATION FOR SEQ ID NO:36:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 28 amino acids
  - (B) TYPE: amino acid
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: protein
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:36:

Met Glu Glu Lys Met Lys Ala Met Glu Glu Lys Met Lys Ala Met Glu  
   1                    5                    10                    15  
 Glu Lys Met Lys Ala Met Glu Glu Lys Met Lys Ala  
                     20                    25

(2) INFORMATION FOR SEQ ID NO:37:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 118 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- (vi) ORIGINAL SOURCE:
  - (B) STRAIN: E. coli
  - (G) CELL TYPE: DH5 alpha
- (vii) IMMEDIATE SOURCE:
  - (B) CLONE: D20
- (ix) FEATURE:
  - (A) NAME/KEY: CDS
  - (B) LOCATION: 2..109
  - (D) OTHER INFORMATION: /function= "synthetic storage protein"
  - /product= "protein"
  - /gene= "ssp"
  - /standard\_name= "5.5.5.5.5"
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:37:

C ATG GAG GAG AAG ATG AAG GCG ATG GAG GAG AAG ATG AAG GCG ATG 46  
 Met Glu Glu Lys Met Lys Ala Met Glu Glu Lys Met Lys Ala Met  
   1                    5                    10                    15  
 GAG GAG AAG ATG AAG GCG ATG GAG GAG AAG ATG AAG GCG ATG GAA GAG 94  
 Glu Glu Lys Met Lys Ala Met Glu Glu Lys Met Lys Ala Met Glu Glu  
                     20                    25                    30

AAG ATG AAG GCG TGATAGGTAC CG  
Lys Met Lys Ala  
35

118

(2) INFORMATION FOR SEQ ID NO:38:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 35 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:38:

Met Glu Glu Lys Met Lys Ala Met Glu Glu Lys Met Lys Ala Met Glu  
1 5 10 15

Glu Lys Met Lys Ala Met Glu Glu Lys Met Lys Ala Met Glu Glu Lys  
20 25 30

Met Lys Ala  
35

(2) INFORMATION FOR SEQ ID NO:39:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 97 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:

- (B) STRAIN: E. coli
- (G) CELL TYPE: DH5 alpha

(vii) IMMEDIATE SOURCE:

- (B) CLONE: D33

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 2..88
- (D) OTHER INFORMATION: /function= "synthetic storage protein"  
/product= "protein"  
/gene= "ssp"  
/standard\_name= "5.5.5.5"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:39:

C ATG GAG GAG AAG ATG AAG GCG ATG GAG GAG AAG ATG AAG GCG ATG 46  
Met Glu Glu Lys Met Lys Ala Met Glu Glu Lys Met Lys Ala Met  
1 5 10 15

GAG GAG AAG ATG AAG GCG ATG GAA GAG AAG ATG AAG GCG TGATAGGTAC 95  
Glu Glu Lys Met Lys Ala Met Glu Glu Lys Met Lys Ala  
20 25

(2) INFORMATION FOR SEO ID NO:40:

(ii) MOLECULE TYPE: protein

Met Glu Glu Lys Met Lys Ala Met Glu Glu Lys Met Lys Ala Met Glu  
1 5 10 15

(2) INFORMATION FOR SEQ ID NO:41:

(xi) SEQUENCE DESCRIPTION: SEO ID NO:41:

(2) INFORMATION FOR SEQ ID NO:42:

oligonucleotide"  
/standard\_name= "SM  
87"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:42:

ATCTTCTTCA GCTTCTCCTC C

21

(2) INFORMATION FOR SEQ ID NO:43:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 21 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

- (A) NAME/KEY: misc\_feature
- (B) LOCATION: 1..21
- (D) OTHER INFORMATION: /product= "synthetic  
oligonucleotide"  
/standard\_name= "SM  
88"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:43:

GATGGAGGAG AAGCTGAAGT G

21

(2) INFORMATION FOR SEQ ID NO:44:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 21 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

- (A) NAME/KEY: misc\_feature
- (B) LOCATION: 1..21
- (D) OTHER INFORMATION: /product= "synthetic  
oligonucleotide"  
/standard\_name= "SM  
89"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:44:

ATCCACTTCA GCTTCTCCTC C

21

(2) INFORMATION FOR SEQ ID NO:45:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 21 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

(A) NAME/KEY: misc\_feature

(B) LOCATION: 1..21

(D) OTHER INFORMATION: /product= "synthetic  
oligonucleotide"  
/standard\_name= "SM  
90"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:45:

GATGGAGGAG AAGATGAAGA A

21

(2) INFORMATION FOR SEQ ID NO:46:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 21 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

(A) NAME/KEY: misc\_feature

(B) LOCATION: 1..21

(D) OTHER INFORMATION: /product= "synthetic  
oligonucleotide"  
/standard\_name= "SM  
91"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:46:

ATCTTCTTCA TCTTCTCCTC C

21

(2) INFORMATION FOR SEQ ID NO:47:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 21 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

(A) NAME/KEY: misc\_feature

(B) LOCATION: 1..21

(D) OTHER INFORMATION: /product= "synthetic  
oligonucleotide"  
/standard\_name= "SM  
92"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:47:

GATGGAGGAG AAGATGAAGT G

21

(2) INFORMATION FOR SEQ ID NO:48:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 21 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- (ix) FEATURE:
  - (A) NAME/KEY: misc\_feature
  - (B) LOCATION: 1..21
  - (D) OTHER INFORMATION: /product= "synthetic oligonucleotide"
  - /standard\_name= "SM 93"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:48:

ATCCACTTCA TCTTCTCCTC C

21

(2) INFORMATION FOR SEQ ID NO:49:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 7 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS: unknown
  - (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:49:

Met Glu Glu Lys Leu Lys Lys  
1 5

(2) INFORMATION FOR SEQ ID NO:50:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 7 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS: unknown
  - (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:50:

Met Glu Glu Lys Leu Lys Trp  
1 5

(2) INFORMATION FOR SEQ ID NO:51:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 7 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS: unknown
  - (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:51:

Met Glu Glu Lys Met Lys Lys  
1 5

(2) INFORMATION FOR SEQ ID NO:52:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 7 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: unknown

(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:52:

Met Glu Glu Lys Met Lys Trp  
1 5

(2) INFORMATION FOR SEQ ID NO:53:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 160 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:

(B) STRAIN: E. coli

(G) CELL TYPE: DH5 alpha

(vii) IMMEDIATE SOURCE:

(B) CLONE: 82-4

(ix) FEATURE:

(A) NAME/KEY: CDS

(B) LOCATION: 2..151

(D) OTHER INFORMATION: /function= "synthetic  
storage protein

/product= "protein"

/gene= "ssp"

/standard\_name=

"7.7.7.7.7.5"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:53:

C ATG GAG GAG AAG CTG AAG GCG ATG GAG GAG AAG CTG AAG GCG ATG 46  
Met Glu Glu Lys Leu Lys Ala Met Glu Glu Lys Leu Lys Ala Met  
1 5 10 15

GAG GAG AAG CTG AAG GCG ATG GAG GAG AAG CTG AAG GCG ATG GAG GAG 94  
Glu Glu Lys Leu Lys Ala Met Glu Glu Lys Leu Lys Ala Met Glu Glu  
20 25 30

AAG CTG AAG GCG ATG GAG GAG AAG CTG AAG GCG ATG GAA GAG AAG ATG 142  
Lys Leu Lys Ala Met Glu Glu Lys Leu Lys Ala Met Glu Glu Lys Met  
35 40 45

AAG GCG TGATAGGTAC CG  
Lys Ala

160

50

(2) INFORMATION FOR SEQ ID NO:54:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 49 amino acids
  - (B) TYPE: amino acid
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:54:

Met Glu Glu Lys Leu Lys Ala Met Glu Glu Lys Leu Lys Ala Met Glu  
1 5 10 15  
Glu Lys Leu Lys Ala Met Glu Glu Lys Leu Lys Ala Met Glu Glu Lys  
20 25 30  
Leu Lys Ala Met Glu Glu Lys Leu Lys Ala Met Glu Glu Lys Met Lys  
35 40 45

Ala

(2) INFORMATION FOR SEQ ID NO:55:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 97 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:

- (B) STRAIN: E. coli
- (G) CELL TYPE: DH5 alpha

(vii) IMMEDIATE SOURCE:

- (B) CLONE: 84-H3

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 2..88
- (D) OTHER INFORMATION: /function= "synthetic  
storage protein  
/product= "protein"  
/gene= "ssp"  
/standard\_name=  
"5.5.5.5"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:55:

C ATG GAG GAG AAG ATG AAG GCG ATG GAG GAG AAG ATG AAG GCG ATG 46  
Met Glu Glu Lys Met Lys Ala Met Glu Glu Lys Met Lys Ala Met  
1 5 10 15





## (2) INFORMATION FOR SEQ ID NO:58:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 28 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:58:

```

Met Glu Glu Lys Met Lys Ala Met Glu Glu Lys Leu Lys Lys Met Glu
 1             5             10             15
Glu Lys Leu Lys Lys Met Glu Glu Lys Met Lys Ala
      20             25

```

## (2) INFORMATION FOR SEQ ID NO:59:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 112 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA (genomic)

## (vi) ORIGINAL SOURCE:

- (B) STRAIN: E. coli
- (G) CELL TYPE: DH5 alpha

## (vii) IMMEDIATE SOURCE:

- (B) CLONE: 88-2

## (ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 2..103
- (D) OTHER INFORMATION: /function= "synthetic  
storage protein  
/product= "protein"  
/gene= "ssp"  
/standard\_name=  
"5.9.9.9.5"

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:59:

```

C ATG GAG GAG AAG ATG AAG GCG AAG AAG CTG AAG TGG ATG GAG GAG   46
  Met Glu Glu Lys Met Lys Ala Lys Lys Leu Lys Trp Met Glu Glu
   1             5             10             15
AAG CTG AAG TGG ATG GAG GAG AAG CTG AAG TGG ATG GAA GAG AAG ATG  94
Lys Leu Lys Trp Met Glu Glu Lys Leu Lys Trp Met Glu Glu Lys Met
      20             25             30
AAG GCG TGATAGGTAC CG                                     112
Lys Ala

```

(2) INFORMATION FOR SEQ ID NO:60:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 33 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:60:

Met Glu Glu Lys Met Lys Ala Lys Lys Leu Lys Trp Met Glu Glu Lys  
1 5 10 15

Leu Lys Trp Met Glu Glu Lys Leu Lys Trp Met Glu Glu Lys Met Lys  
20 25 30

Ala

(2) INFORMATION FOR SEQ ID NO:61:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 118 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:

- (B) STRAIN: E. coli
- (G) CELL TYPE: DH5 alpha

(vii) IMMEDIATE SOURCE:

- (B) CLONE: 90-H8

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 2..109
- (D) OTHER INFORMATION: /function= "synthetic storage protein  
/product= "protein"  
/gene= "ssp"  
/standard\_name=  
"5.10.10.10.5"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:61:

C ATG GAG GAG AAG ATG AAG GCG ATG GAG GAG AAG ATG AAG AAG ATG 46  
Met Glu Glu Lys Met Lys Ala Met Glu Glu Lys Met Lys Lys Met  
1 5 10 15

GAG GAG AAG ATG AAG AAG ATG GAG GAG AAG ATG AAG AAG ATG GAA GAG 94  
Glu Glu Lys Met Lys Lys Met Glu Glu Lys Met Lys Lys Met Glu Glu  
20 25 30

AAG ATG AAG GCG TGATAGGTAC CG 118  
Lys Met Lys Ala  
35

(2) INFORMATION FOR SEQ ID NO:62:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 35 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:62:

```
Met Glu Glu Lys Met Lys Ala Met Glu Glu Lys Met Lys Lys Met Glu
 1             5             10             15
Glu Lys Met Lys Lys Met Glu Glu Lys Met Lys Lys Met Glu Glu Lys
          20             25             30
Met Lys Ala
          35
```

(2) INFORMATION FOR SEQ ID NO:63:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 97 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:

- (B) STRAIN: E. coli
- (G) CELL TYPE: DH5 alpha

(vii) IMMEDIATE SOURCE:

- (B) CLONE: 92-2

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 2..88
- (D) OTHER INFORMATION: /function= "synthetic  
storage protein  
/product= "protein"  
/gene= "ssp"  
/standard\_name=  
"5.11.11.5"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:63:

```
C ATG GAG GAG AAG ATG AAG GCG ATG GAG GAG AAG ATG AAG TGG ATG      46
  Met Glu Glu Lys Met Lys Ala Met Glu Glu Lys Met Lys Trp Met
  1             5             10             15
GAG GAG AAG ATG AAG TGG ATG GAA GAG AAG ATG AAG GCG TGATAGGTAC    95
Glu Glu Lys Met Lys Trp Met Glu Glu Lys Met Lys Ala
          20             25
CG                                                     97
```

(2) INFORMATION FOR SEQ ID NO:64:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 28 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:64:

```
Met Glu Glu Lys Met Lys Ala Met Glu Glu Lys Met Lys Trp Met Glu
 1             5             10             15
Glu Lys Met Lys Trp Met Glu Glu Lys Met Lys Ala
          20             25
```

(2) INFORMATION FOR SEQ ID NO:65:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 84 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

- (A) NAME/KEY: misc\_feature
- (B) LOCATION: 1..84
- (D) OTHER INFORMATION: /product= "synthetic  
oligonucleotide"  
/standard\_name= "SM  
96"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:65:

```
GATGGAGGAA AAGATGAAGG CGATGGAGGA GAAAATGAAA GCTATGGAGG AAAAGATGAA 60
AGCGATGGAG GAGAAAATGA AGGC                                     84
```

(2) INFORMATION FOR SEQ ID NO:66:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 84 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

- (A) NAME/KEY: misc\_feature
- (B) LOCATION: 1..84
- (D) OTHER INFORMATION: /product= "synthetic  
oligonucleotide"  
/standard\_name= "SM  
97"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:66:

ATCGCCTTCA TTTTCTCCTC CATCGCTTTC ATCTTTTCCT CCATAGCTTT CATTTTCTCC 60  
TCCATCGCCT TCATCTTTTC CTCC 84

(2) INFORMATION FOR SEQ ID NO:67:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 28 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS: unknown
  - (D) TOPOLOGY: unknown
- (ii) MOLECULE TYPE: protein
- (ix) FEATURE:
  - (A) NAME/KEY: Protein
  - (B) LOCATION: 1..28
  - (D) OTHER INFORMATION: /label= name  
/note= "(SSP 5)4"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:67:

Met Glu Glu Lys Met Lys Ala Met Glu Glu Lys Met Lys Ala Met Glu  
1 5 10 15  
Glu Lys Met Lys Ala Met Glu Glu Lys Met Lys Ala  
20 25

(2) INFORMATION FOR SEQ ID NO:68:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 84 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- (ix) FEATURE:
  - (A) NAME/KEY: misc\_feature
  - (B) LOCATION: 1..84
  - (D) OTHER INFORMATION: /product= "synthetic  
oligonucleotide"  
/standard\_name= "SM  
98"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:68:

GATGGAGGAA AAGCTGAAAG CGATGGAGGA GAAACTCAAG GCTATGGAAG AAAAGCTTAA 60  
AGCGATGGAG GAGAAACTGA AGGC 84

(2) INFORMATION FOR SEQ ID NO:69:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 84 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

(A) NAME/KEY: misc\_feature

(B) LOCATION: 1..84

(D) OTHER INFORMATION: /product= "synthetic  
oligonucleotide"  
/standard\_name= "SM  
99"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:69:

ATCGCCTTCA GTTTCCTCCTC CTACGCTTTA AGCTTTTCTT CCATAGCCTT GAGTTTCTCC 60  
TCCATCGCTT TCAGCTTTTC CTCC 84

(2) INFORMATION FOR SEQ ID NO:70:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 28 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: unknown

(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(ix) FEATURE:

(A) NAME/KEY: Protein

(B) LOCATION: 1..28

(D) OTHER INFORMATION: /label= name  
/note= "(SSP 7)4"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:70:

Met Glu Glu Lys Leu Lys Ala Met Glu Glu Lys Leu Lys Ala Met Glu  
1 5 10 15  
Glu Lys Leu Lys Ala Met Glu Glu Lys Leu Lys Ala  
20 25

(2) INFORMATION FOR SEQ ID NO:71:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 84 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

(A) NAME/KEY: misc\_feature

(B) LOCATION: 1..84

(D) OTHER INFORMATION: /product= "synthetic  
oligonucleotide"  
/standard\_name= "SM  
100"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:71:

GATGGAGGAA AAGCTTAAGA AGATGGAAGA AAAGCTGAAA TGGATGGAGG AGAAACTCAA 60  
AAAGATGGAG GAAAAGCTTA AATG 84

(2) INFORMATION FOR SEQ ID NO:72:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 84 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

(A) NAME/KEY: misc\_feature  
(B) LOCATION: 1..84  
(D) OTHER INFORMATION: /product= "synthetic  
oligonucleotide"  
/standard\_name= "SM  
101"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:72:

ATCCATTTAA GCTTTTCCTC CTACTTTTGT AGTTTCTCCT CCATCCATTT CAGCTTTTCT 60  
TCCATCTTCT TAAGCTTTTC CTCC 84

(2) INFORMATION FOR SEQ ID NO:73:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 28 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS: unknown  
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:73:

Met Glu Glu Lys Leu Lys Lys Met Glu Glu Lys Leu Lys Trp Met Glu  
1                      5                      10                      15  
Glu Lys Leu Lys Lys Met Glu Glu Lys Leu Lys Trp  
20                      25

(2) INFORMATION FOR SEQ ID NO:74:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 243 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:

(B) STRAIN: E. coli



(G) CELL TYPE: DH5 alpha

(vii) IMMEDIATE SOURCE:

(B) CLONE: 2-9

(ix) FEATURE:

(A) NAME/KEY: CDS

(B) LOCATION: 2..235

(D) OTHER INFORMATION: /function= "synthetic  
storage protein  
/product= "protein"  
/gene= "ssp"  
/standard\_name=  
"7.7.7.7.7.7.8.9.8.9.5"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:74:

```
C  ATG  GAG  GAG  AAG  CTG  AAG  GCG  ATG  GAG  GAG  AAG  CTG  AAG  GCG  ATG      46
   Met  Glu  Glu  Lys  Leu  Lys  Ala  Met  Glu  Glu  Lys  Leu  Lys  Ala  Met
   1              5              10              15

GAG  GAG  AAG  CTG  AAG  GCG  ATG  GAG  GAG  AAG  CTG  AAG  GCG  ATG  GAG  GAG      94
Glu  Glu  Lys  Leu  Lys  Ala  Met  Glu  Glu  Lys  Leu  Lys  Ala  Met  Glu  Glu
              20              25              30

AAG  CTG  AAG  GCG  ATG  GAG  GAG  AAG  CTG  AAG  GCG  ATG  GAG  GAA  AAG  CTT      142
Lys  Leu  Lys  Ala  Met  Glu  Glu  Lys  Leu  Lys  Ala  Met  Glu  Glu  Lys  Leu
              35              40              45

AAG  AAG  ATG  GAA  GAA  AAG  CTG  AAA  TGG  ATG  GAG  GAG  AAA  CTC  AAA  AAG      190
Lys  Lys  Met  Glu  Glu  Lys  Leu  Lys  Trp  Met  Glu  Glu  Lys  Leu  Lys  Lys
              50              55              60

ATG  GAG  GAA  AAG  CTT  AAA  TGG  ATG  GAA  GAG  AAG  ATG  AAG  GCG  TGATAGGTAC  242
Met  Glu  Glu  Lys  Leu  Lys  Trp  Met  Glu  Glu  Lys  Met  Lys  Ala
   65              70              75

C                                          243
```

(2) INFORMATION FOR SEQ ID NO:75:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 77 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:75:

```
Met  Glu  Glu  Lys  Leu  Lys  Ala  Met  Glu  Glu  Lys  Leu  Lys  Ala  Met  Glu
   1              5              10              15

Glu  Lys  Leu  Lys  Ala  Met  Glu  Glu  Lys  Leu  Lys  Ala  Met  Glu  Glu  Lys
              20              25              30

Leu  Lys  Ala  Met  Glu  Glu  Lys  Leu  Lys  Ala  Met  Glu  Glu  Lys  Leu  Lys
              35              40              45
```

Lys Met Glu Glu Lys Leu Lys Trp Met Glu Glu Lys Leu Lys Lys Met  
50 55 60

Glu Glu Lys Leu Lys Trp Met Glu Glu Lys Met Lys Ala  
65 70 75

(2) INFORMATION FOR SEQ ID NO:76:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 175 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:

- (B) STRAIN: E. coli
- (G) CELL TYPE: DH5 alpha

(vii) IMMEDIATE SOURCE:

- (B) CLONE: 5-1

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 2..172
- (D) OTHER INFORMATION: /function= "synthetic storage protein  
/product= "protein"  
/gene= "ssp"  
/standard\_name= "5.5.5.7.7.7.7.5"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:76:

C ATG GAG GAG AAG ATG AAG GCG ATG GAG GAG AAG ATG AAG GCG ATG	46
Met Glu Glu Lys Met Lys Ala Met Glu Glu Lys Met Lys Ala Met	
1 5 10 15	
GAG GAG AAG ATG AAG GCG ATG GAG GAA AAG CTG AAA GCG ATG GAG GAG	94
Glu Glu Lys Met Lys Ala Met Glu Glu Lys Leu Lys Ala Met Glu Glu	
20 25 30	
AAA CTC AAG GCT ATG GAA GAA AAG CTT AAA GCG ATG GAG GAG AAA CTG	142
Lys Leu Lys Ala Met Glu Glu Lys Leu Lys Ala Met Glu Glu Lys Leu	
35 40 45	
AAG GCC ATG GAA GAG AAG ATG AAG GCG TGATAG	175
Lys Ala Met Glu Glu Lys Met Lys Ala	
50 55	

(2) INFORMATION FOR SEQ ID NO:77:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 56 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:77:

Met Glu Glu Lys Met Lys Ala Met Glu Glu Lys Met Lys Ala Met Glu  
1 5 10 15  
Glu Lys Met Lys Ala Met Glu Glu Lys Leu Lys Ala Met Glu Glu Lys  
20 25 30  
Leu Lys Ala Met Glu Glu Lys Leu Lys Ala Met Glu Glu Lys Leu Lys  
35 40 45  
Ala Met Glu Glu Lys Met Lys Ala  
50 55

(2) INFORMATION FOR SEQ ID NO:78:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 187 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:

- (B) STRAIN: E. coli
- (G) CELL TYPE: DH5 alpha

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 3..173
- (D) OTHER INFORMATION: /function= "synthetic storage protein  
/product= "protein"  
/gene= "ssp"  
/standard\_name=  
"SSP-3-5"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:78:

CC ATG GAG GAG AAG CTG AAG GCG ATG GAG GAG AAG CTG AAG GCG ATG 47  
Met Glu Glu Lys Leu Lys Ala Met Glu Glu Lys Leu Lys Ala Met  
1 5 10 15  
GAG GAG AAG CTG AAG GCG ATG GAG GAG AAG CTG AAG GCG ATG GAG GAG 95  
Glu Glu Lys Leu Lys Ala Met Glu Glu Lys Leu Lys Ala Met Glu Glu  
20 25 30  
AAG CTG AAG GCG ATG GAG GAG AAG CTG AAG GCG ATG GAG GAA AAG ATG 143  
Lys Leu Lys Ala Met Glu Glu Lys Leu Lys Ala Met Glu Glu Lys Met  
35 40 45  
AAG GCG ATG GAA GAG AAG ATG AAG GCG TGATAGGTAC CGAATTC 187  
Lys Ala Met Glu Glu Lys Met Lys Ala  
50 55

(2) INFORMATION FOR SEQ ID NO:79:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 56 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:79:

```
Met Glu Glu Lys Leu Lys Ala Met Glu Glu Lys Leu Lys Ala Met Glu
 1           5           10           15
Glu Lys Leu Lys Ala Met Glu Glu Lys Leu Lys Ala Met Glu Glu Lys
          20           25           30
Leu Lys Ala Met Glu Glu Lys Leu Lys Ala Met Glu Glu Lys Met Lys
          35           40           45
Ala Met Glu Glu Lys Met Lys Ala
 50           55
```

(2) INFORMATION FOR SEQ ID NO:80:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 61 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

(A) NAME/KEY: misc\_feature

(B) LOCATION: 1..61

(D) OTHER INFORMATION: /product= "synthetic  
oligonucleotide"  
/standard\_name= "SM  
107"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:80:

```
CATGGAGGAG AAGATGAAAA AGCTCGAAGA GAAGATGAAG GTCATGAAGT GATAGGTACC 60
G                                                                 61
```

(2) INFORMATION FOR SEQ ID NO:81:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 61 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

(A) NAME/KEY: misc\_feature

(B) LOCATION: 1..61

(D) OTHER INFORMATION: /product= "synthetic  
ligonucleotide"  
/standard\_name= "SM  
106"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:81:

AATTCGGTAC CTATCACTTC ATGACCTTCA TCTTCTCTTC GAGCTTTTTC ATCTTCTCCT 60

C

61

(2) INFORMATION FOR SEQ ID NO:82:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 16 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: unknown

(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(ix) FEATURE:

(A) NAME/KEY: Protein

(B) LOCATION: 1..16

(D) OTHER INFORMATION: /label= name

/note= "pSK34 base

gene"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:82:

Met Glu Glu Lys Met Lys Lys Leu Glu Glu Lys Met Lys Val Met Lys  
1 5 10 15

(2) INFORMATION FOR SEQ ID NO:83:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 63 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

(A) NAME/KEY: misc\_feature

(B) LOCATION: 1..63

(D) OTHER INFORMATION: /product= "synthetic  
oligonucleotide"

/standard\_name= "SM

110"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:83:

GCTGGAAGAA AAGATGAAGG CTATGGAGGA CAAGATGAAA TGGCTTGAGG AAAAGATGAA 60

GAA

63

(2) INFORMATION FOR SEQ ID NO:84:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 63 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

(A) NAME/KEY: misc\_feature

(B) LOCATION: 1..63

(D) OTHER INFORMATION: /product= "synthetic  
oligonucleotide"  
/standard\_name= "SM  
111"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:84:

AGCTTCTTCA TCTTTTCCTC AAGCCATTTT ATCTTGTCCT CCATAGCCTT CATCTTTTCT 60  
TCC 63

(2) INFORMATION FOR SEQ ID NO:85:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 37 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:85:

Met Glu Glu Lys Met Lys Lys Leu Glu Glu Lys Met Lys Ala Met Glu  
1 5 10 15  
Asp Lys Met Lys Trp Leu Glu Glu Lys Met Lys Lys Leu Glu Glu Lys  
20 25 30  
Met Lys Val Met Lys  
35

(2) INFORMATION FOR SEQ ID NO:86:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 37 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:86:

Met Glu Glu Lys Met Lys Lys Leu Glu Glu Lys Met Lys Ala Met Glu  
1 5 10 15  
Asp Lys Met Lys Trp Leu Glu Glu Lys Met Lys Lys Leu Glu Glu Lys  
20 25 30  
Met Lys Val Met Lys  
35

(2) INFORMATION FOR SEQ ID NO:87:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 62 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:  
 (A) NAME/KEY: misc\_feature  
 (B) LOCATION: 1..62  
 (D) OTHER INFORMATION: /product= "synthetic  
 oligonucleotide"  
 /standard\_name= "SM  
 112"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:87:

GCTCGAAGAA AGATGAAGGC AATGGAAGAC AAAATGAAGT GGCTTGAGGA GAAAATGAAG 60  
 AA 62

(2) INFORMATION FOR SEQ ID NO:88:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 62 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:  
 (A) NAME/KEY: misc\_feature  
 (B) LOCATION: 1..62  
 (D) OTHER INFORMATION: /product= "synthetic  
 oligonucleotide"  
 /standard\_name= "SM  
 113"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:88:

AGCTTCTTCA TTTTCTCCTC AAGCCACTTC ATTTTGTCTT CCATTGCCTT CATCTTTCTT 60  
 CG 62

(2) INFORMATION FOR SEQ ID NO:89:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 37 amino acids  
 (B) TYPE: amino acid  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:89:

Met Glu Glu Lys Met Lys Lys Leu Lys Glu Glu Met Ala Lys Met Lys  
 1 5 10 15

Asp Glu Met Trp Lys Leu Lys Glu Glu Met Lys Lys Leu Glu Glu Lys  
20 25 30

Met Lys Val Met Lys  
35

(2) INFORMATION FOR SEQ ID NO:90:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 63 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- (ix) FEATURE:
  - (A) NAME/KEY: misc\_feature
  - (B) LOCATION: 1..63
  - (D) OTHER INFORMATION: /product= "synthetic oligonucleotide"
  - /standard\_name= "SM 114"
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:90:

GCTCAAGGAG GAAATGGCTA AGATGAAAGA CGAAATCTGG AAAGTAAAG AGGAAATGAA 60  
GAA 63

(2) INFORMATION FOR SEQ ID NO:91:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 63 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- (ix) FEATURE:
  - (A) NAME/KEY: misc\_feature
  - (B) LOCATION: 1..63
  - (D) OTHER INFORMATION: /product= "synthetic oligonucleotide"
  - /standard\_name= "SM 115"
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:91:

AGCTTCTTCA TTTCCTCTTT CAGTTTCCAC ATTTCTGCTT TCATCTTAGC CATTCCTCC 60  
TTG 63

(2) INFORMATION FOR SEQ ID NO:92:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 107 amino acids
  - (B) TYPE: amino acid
  - (D) TOPOLOGY: linear



(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:92:

```
Met Glu Glu Lys Met Lys Lys Leu Lys Glu Glu Met Ala Lys Met Lys
 1           5           10           15
Asp Glu Met Trp Lys Leu Lys Glu Glu Met Lys Lys Leu Glu Glu Lys
 20           25           30
Met Lys Val Met Glu Glu Lys Met Lys Lys Leu Glu Glu Lys Met Lys
 35           40           45
Ala Met Glu Asp Lys Met Lys Trp Leu Glu Glu Lys Met Lys Lys Leu
 50           55           60
Glu Glu Lys Met Lys Val Met Glu Glu Lys Met Lys Lys Leu Glu Glu
 65           70           75           80
Lys Met Lys Ala Met Glu Asp Lys Met Lys Trp Leu Glu Glu Lys Met
 85           90           95
Lys Lys Leu Glu Glu Lys Met Lys Val Met Lys
 100           105
```

(2) INFORMATION FOR SEQ ID NO:93:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 839 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:93:

```
GGATCCCCCG GGCTGCAGGA ATTCTACGTA CCATATAGTA AGACTTTGTA TATAAGACGT 60
CACCTCTTAC GTGCATGGTT ATATGTGACA TGTGCAGTGA CGTTGTACCA TATAGTAAGA 120
CTTTGTATAT AAGACGTCAC CTCTTACGTG CATGGTTATA TGTGACATGT GCAGTGACGT 180
TAACCGCACC CTCCTTCCCG TCGTTTCCCA TCTCTTCCTC CTTTAGAGCT ACCACTATAT 240
AAATCAGGGC TCATTTTCTC GCTCCTCACA GGCTCATCAG CACCCCGGCA GTGCCACCCC 300
GACTCCCTGC ACCTGCCATG GGTACGCTAG CCCGGGAGAT CTGACAAAGC AGCATTAGTC 360
CGTTGATCGG TGGAAGACCA CTCGTCAGTG TTGAGTTGAA TGTTTGATCA ATAAAATACG 420
GCAATGCTGT AAGGGTTGTT TTTTATGCCA TTGATAATAC ACTGTACTGT TCAGTTGTTG 480
AACTCTATTT CTTAGCCATG CCAGTGCTTT TCTTATTTTG AATAACATTA CAGCAAAAAG 540
TTGAAAGACA AAAAAANNNN NCCCCGAACA GAGTGCTTTG GGTCCCAAGC TTCTTTAGAC 600
TGTGTTTCGGC GTTCCCCCTA AATTTCTCCC CTATATCTCA CTCACTTGTC ACATCAGCGT 660
```

TCTCTTTCCC CTATATCTCC ACGCTCTACA GCAGTTCCAC CTATATCAAA CCTCTATACC 720  
 CCACCACAAC AATATTATAT ACTTTCATCT TCACCTAACT CATGTACCTT CCAATTTTTT 780  
 TCTACTAATA ATTATTTACG TGCACAGAAA CTTAGGCAAG GGAGAGAGAG AGCGGTACC 839

(2) INFORMATION FOR SEQ ID NO:94:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 43 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:94:

CTAGAAGCCT CGGCAACGTC AGCAACGGCG GAAGAATCCG GTG 43

(2) INFORMATION FOR SEQ ID NO:95:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 43 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:95:

CATGCACCGG ATTCTTCCGC CGTTGCTGAC GTTGCCGAGG CTT 43

(2) INFORMATION FOR SEQ ID NO:96:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 55 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:96:

GATCCCATGG CGCCCCCTTAA GTCCACCGCC AGCCTCCCCG TCGCCCGCCG CTCCT 55

(2) INFORMATION FOR SEQ ID NO:97:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 55 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:97:

CTAGAGGAGC GCGGGGCGAC GGGGAGGCTG GCGGTGGACT TAAGGGGCGC CATGG 55

(2) INFORMATION FOR SEQ ID NO:98:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 59 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:98:

CATGGCGCCC ACCGTGATGA TGGCCTCGTC GGCCACCGCC GTCGCTCCGT TCCAGGGGC 59

(2) INFORMATION FOR SEQ ID NO:99:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 59 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:99:

TTAAGCCCCT GGAACGGAGC GACGGCGGTG GCCGACGAGG CCATCATCAC GGTGGGCGC 59

(2) INFORMATION FOR SEQ ID NO:100:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 16 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:100:

GCGCCACCG TGATGA 16

(2) INFORMATION FOR SEQ ID NO:101:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 16 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:101:

CACCGGATTC TTCCGC 16

(2) INFORMATION FOR SEQ ID NO:102:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 372 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:102:

```
GTAAGATTGG TAAAGTCCAG CAAGAAAATG AGATAAAAGA GAAGCCTGAA ATGACGAAAA    60
AATCAGGTGT TTTGATTCTT GGTGCTGGAC GTGTGTNTCG CCCAGCTGCT GATTTCTTAG    120
CTTCAGTTAG AACCATTTCG TCACAGCAAT GGTACAAAAC ATATTTTCGGA GCAGACTCTG    180
AAGAGAAAAC AGATGTTTCAT GTGATTGTCG CGTCTCTGTA TCTTAAGGAT GCCAAAGAGA    240
CGGTTGAAGG TATTTTCAGAT GTAGAAGCAG TTCGGCTAGA TGTATCTGAT AGTGAAAGTC    300
TCCTTAAGTA TGTTTCTCAG GTTGATGTTG TCCTAAGTTT ATTACCTGCA AGTTGTCATG    360
CTTGTTGTAG CA                                                    372
```

(2) INFORMATION FOR SEQ ID NO:103:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 323 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:103:

```
GGAAGCACAC TGCGACTCTT TTGGAATTCG GGGACATCAA GAATGGACAA ACAACAACCG    60
CTATGGCCAA GACTGTTGGG ATCCCTGCAG CCATTGGAGC TCTGCTGTTA ATTGAAGACA    120
AGATCAAGAC AAGAGGAGTC TTAAGGCCTC TCGAAGCAGA GGTGTATTTG CCAGCTTTGG    180
ATATATTGCA AGCATATGGT ATAAAGCTGA TGGAGAAGGC AGAATGATCA AAGAACTCTG    240
TATATTGTTT CTNTCTATAA CTTGGAGTTG GAGACAAAGC TGAAGGAGNC AGNGCCATTA    300
GACCAGCAAA AAAAGGAGGA GGA                                                    323
```

(2) INFORMATION FOR SEQ ID NO:104:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 123 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:104:

Lys Ile Gly Lys Val Gln Gln Glu Asn Glu Ile Lys Glu Lys Pro Glu  
 1 5 10 15  
 Met Thr Lys Lys Ser Gly Val Leu Ile Leu Gly Ala Gly Arg Val Xaa  
 20 25 30  
 Arg Pro Ala Ala Asp Phe Leu Ala Ser Val Arg Thr Ile Ser Ser Gln  
 35 40 45  
 Gln Trp Tyr Lys Thr Tyr Phe Gly Ala Asp Ser Glu Glu Lys Thr Asp  
 50 55 60  
 Val His Val Ile Val Ala Ser Leu Tyr Leu Lys Asp Ala Lys Glu Thr  
 65 70 75 80  
 Val Glu Gly Ile Ser Asp Val Glu Ala Val Arg Leu Asp Val Ser Asp  
 85 90 95  
 Ser Glu Ser Leu Leu Lys Tyr Val Ser Gln Val Asp Val Val Leu Ser  
 100 105 110  
 Leu Leu Pro Ala Ser Cys His Ala Cys Cys Ser  
 115 120

(2) INFORMATION FOR SEQ ID NO:105:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 74 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:105:

Lys His Thr Ala Thr Leu Leu Glu Phe Gly Asp Ile Lys Asn Gly Gln  
 1 5 10 15  
 Thr Thr Thr Ala Met Ala Lys Thr Val Gly Ile Pro Ala Ala Ile Gly  
 20 25 30  
 Ala Leu Leu Leu Ile Glu Asp Lys Ile Lys Thr Arg Gly Val Leu Arg  
 35 40 45  
 Pro Leu Glu Ala Glu Val Tyr Leu Pro Ala Leu Asp Ile Leu Gln Ala  
 50 55 60  
 Tyr Gly Ile Lys Leu Met Glu Lys Ala Glu  
 65 70

(2) INFORMATION FOR SEQ ID NO:106:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 25 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:106:  
ATTCCCCATG GTTTCGCCGA CGAAT 25

(2) INFORMATION FOR SEQ ID NO:107:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 29 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:107:  
CTCTCGGTAC CTAGTACCTA CTGATCAAC 29

(2) INFORMATION FOR SEQ ID NO:108:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 56 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:108:  
Met Glu Glu Lys Leu Lys Ala Met Glu Glu Lys Leu Lys Ala Met Glu  
1 5 10 15  
Glu Lys Leu Lys Ala Met Glu Glu Lys Leu Lys Ala Met Glu Glu Lys  
20 25 30  
Leu Lys Ala Met Glu Glu Lys Leu Lys Ala Met Glu Glu Lys Met Lys  
35 40 45  
Ala Met Glu Glu Lys Met Lys Ala  
50 55